

Data scientist junior, spécialisé dans le développement de modèles prédictifs pour résoudre des problèmes complexes, notamment l'intégration d'algorithmes d'IA dans le contrôle des processus industriels. Compétences solides en apprentissage automatique, analyse de données, résolution de problèmes et programmation. Maîtrise des technologies Big Data comme Apache Spark, Kafka et Docker, avec une passion pour transformer les données en solutions stratégiques.

EDUCATION

MS Systèmes Informatiques Distribués & Big Data , <i>Faculté des Sciences - Agadir</i>	2022 — 2024
Licence Sciences Mathématiques et Informatique , <i>Faculté Polydisciplinaire de Ouarzazate</i>	2019 — 2022
Baccalauréat Sciences Expérimentales filière Physique , <i>Lycée Othman Ibn Affane Tagounite</i>	2018 — 2019

COMPÉTENCES

Langages	Python, Java, C
Bases de Données	PostgreSQL, MySQL
Frameworks Big Data	Apache Hadoop, Apache Spark, Apache Kafka, Docker, Talend, ETL Processes
Bibliothèques	pandas, NumPy, Sklearn, SciPy, Keras, nltk, matplotlib, TensorFlow, Flask, React JS
Algorithms	Régression, Random Forest, Boosted Decision Trees, SVM, K-means Clustering, PCA, Deep Learning Algorithms, Analyse de Séries Temporelles
Deep Learning	NN, RNN, LSTM, CNN, Optimization of Algorithms, Regularization, Hyperparameter Tuning
Statistiques	Statistiques Descriptives, Inférence Statistique

EXPÉRIENCES TECHNIQUES

Data Scientist

Conception et implémentation d'un algorithme de contrôle MPC pour l'optimisation des procédés industriels dans la production d'acide phosphorique

Teal Technology Services

03/2024 — 09/2024
MAROC, CASABLANCA

- Nettoyage et prétraitement des données.
- Gestion des valeurs aberrantes (outliers).
- Analyse et extraction des tendances.
- Application d'algorithmes tels que la régression linéaire, les réseaux de neurones, les arbres de décision, les forêts aléatoires, XGBoost, et évaluation de leurs performances (MAE, MSE, R-square).
- Sélection du meilleur algorithme (NN) parmi eux et intégration dans le contrôle prédictif par modèle (MPC)

PROJECTS

Analyse des données aériens en temps réel

- Collecte de données en temps réel (API)
- ingestion des données (kafka)
- Nettoyage et prétraitement des données (Pyspark)
- Création de tableaux de bord avec Kibana et Elasticsearch, offrant des informations en temps réel pour une meilleure prise de décision.

Analyse en temps réel des données criminelles à Boston.

- Récupérer des données de l'API de la police de Boston
- Nettoyage et prétraitement des données (Pyspark)
- Développer un modèle LSTM pour prédire les crimes en fonction de facteurs clés, tels que la zone géographique et le nombre de déclarations de crimes dans cette zone.

Analyse des accidents aériens.

- Le jeu de données utilisé pour l'analyse provient de Kaggle et regroupe tous les accidents survenus de 1908 à 2009.
- Nettoyage et prétraitement des données
- Analyser les données pour détecter des motifs et des tendances dans les accidents.
- Visualisation des résultats avec Plotly et Matplotlib

CERTIFICATS

Machine Learning Specialization - Coursera

[Lien](#)

The Machine Learning Algorithms A-Z - 365datascience

[Lien](#)

Python for Data Science, AI & Development - IBM

[Lien](#)

Hadoop: Map Reduce - Great Learning

[Lien](#)

Getting Started with Git and GitHub - IBM

[Lien](#)